PLOS ONE

RESEARCH ARTICLE

Bias of marker genes in PCR of anammox bacteria in natural habitats

Min Cai, Fei Ye, Jiapeng Wu, Qihang Wu, Yu Wang⬦*◔**, Yiguo Hong***◔

Institute of Environmental Research at Greater Bay Area, Key Laboratory for Water Quality and Conservation of the Pearl River Delta, Ministry of Education, Guangzhou University, Guangzhou, China

◔ These authors contributed equally to this work.
* wangyu@gzhu.edu.cn (YW); yghong@gzhu.edu.cn (YH)

Abstract

The identification of anammox bacteria is mostly relied on PCR with various marker genes. However, the community composition revealed by different marker genes and whether the marker genes influence the resulted community composition remain unclear. We compared the community structure of anammox bacteria in enriched and natural environments revealed by 16S rRNA and functional genes (*hzo*, *hzsA* and *hzsB*) from public database and published papers. The genus of *Ca. Scalindua* showed the lowest similarities with other genera, especially for the *hzsA* gene (66.9%-68.6%). The 16S rRNA gene is the most commonly used marker gene in natural habitats with 151 out 221 papers in total. The anammox bacterial community composition is distributed according to the source of habitat regardless the use of various marker genes. The role of marker gene is limited with explanatory of 5.4% for variance of community composition, versus 20.5% of habitat. The effect of marker gene is mainly acted on freshwater habitat, which shows significant different community composition revealed by 16S rRNA and *hzo*, with *Ca. Brocadia* and *Ca. Jettenia* as dominant genus, respectively.

Introduction

Anammox (anaerobic ammonium oxidation) is the most important finding in nitrogen cycle for the last two decades [1]. Anammox bacteria deeply branch in *Planctomycetales*, and contribute greatly in natural habitats (responsible for 50%, 9~40%, and 4~37% of the nitrogen loss for marine, lakes, and paddy soils) [2–4]. There are five genera of anammox bacteria were identified to date including Ca. *Brocadia* [5], Ca. *Kuenenia* [6], Ca. *Scalindua* [7], Ca. *Anammoxoglobus* [8], and Ca. *Jettenia* [9]. Each genus shows distinct physiological property and adapts to specific habitat [10]. Hence, the community composition of anammox bacteria is very diverse in natural habitats.

The detection of anammox bacteria is greatly relied on PCR based amplification which is an easy, essential, and widely accessible technique [11]. The 16S rRNA gene is the most used marker gene offering phylogeny details in various researches. Besides, many functional genes targeting central anammox activity enzymes were applied for the amplification of anammox bacteria. The functional genes were less used but have high specificity and reflected functional

activity [12]. The PCR bias by different maker genes often skews the distribution of PCR products due to the unequal amplification or cloning efficiency [13]. The choosing of marker genes was empiric and varied largely in different habitats in PCR of anammox bacteria. However, how is the community composed revealed by various marker genes, and whether the usage of different marker genes would affect the resulted anammox bacterial community remained unclear. Here we compared the enriched anammox bacteria revealed by different marker genes to show their similarities on genus level. Moreover, the community composition in natural habitats was summarized and compared from published peer-review publications to show the influence of marker genes.

Methods and materials

The representative sequences of enriched anammox bacteria of each genus was retrieved from GenBank. The similarities of the 16S rRNA and functional genes between genera were obtained by blastx in NCBI and compared pairwisely. Peer-reviewed anammox related papers in natural habitats were searched in Web of Science (published in 2006.10–2020.04) with keywords of *anammox* and *anaerobic ammonium oxidation*. The community composition was summarized according to their original description in each study. The diversity indexes (Shannon Simpson Chao 1 and Ace) were calculated by Mothur program using the collected sequences in each paper (v.1.40.1, https://www.mothur.org/wiki/).

Non-metric multidimensional scaling (NMDS) based on Bray-Curtis dissimilarity was conducted in CANOCO 5 [14]. Similarly, analysis of similarity (ANOSIM) based on Bray-Curtis distances was conducted on community structure difference test using R (version 3.5.1, vegan package). Variation partitioning analysis (VPA) was then used to quantify the relative influences of habitat and marker gene on the anammox bacterial community in CANOCO 5 [14].

Results

Comparison of enriched anammox bacteria

Representative nucleotide sequences for the 5 known genera of the frequently used genes: 16S rRNA and functional genes (*hzo*, *hzsA* and *hzsB*) were obtained from Genbank (S1 Table). Sequence of Ca. *Anammoxoglobus* was missing in *hzsB* gene. The nucleotide identities (NI) between the representative sequences of the 5 known genera varied considerably between 16S rRNA and functional genes. The NI of 16S rRNA gene showed the highest values with 90.0% to 95.6% identities, and the *hzsA* gene showed the lowest similarities with NI of 66.9% to 85.7% (Fig 1A). The genus of Ca. *Scalindua* showed the largest difference with other 4 genera in all genes, in which the NI were 89.0%-90.3% for 16S rRNA gene, 79.5%-81.1% for *hzo* gene, 66.9%-68.6% for *hzsA* gene and 81.2%-83.3% for *hzsB* gene. Aside from Ca. *Scalindua*, the other 4 genera showed relatively high similarity with NI of 92.9%-95.6% for 16S rRNA, 81.6%-89.4% for *hzo*, 83.7%-85.7% for *hzsA* and 81.2–87.0% for *hzsB*. The dissimilarities between genera revealed by functional and 16S rRNA genes varied largely except the *hzo* and 16S rRNA genes (Fig 2B).

Community composition in natural habitats

According to the habitat preference of anammox bacteria, the 221 collected studies were divided into four groups which were marine, estuary, freshwater and soil environments (Table 1). The freshwater environment contained a wide range of habitats including lake, river and freshwater wetland. The 16S rRNA gene was the most used marker gene studying the anammox bacteria which included 151 out 221 studies. The *hzo* gene was the second most used
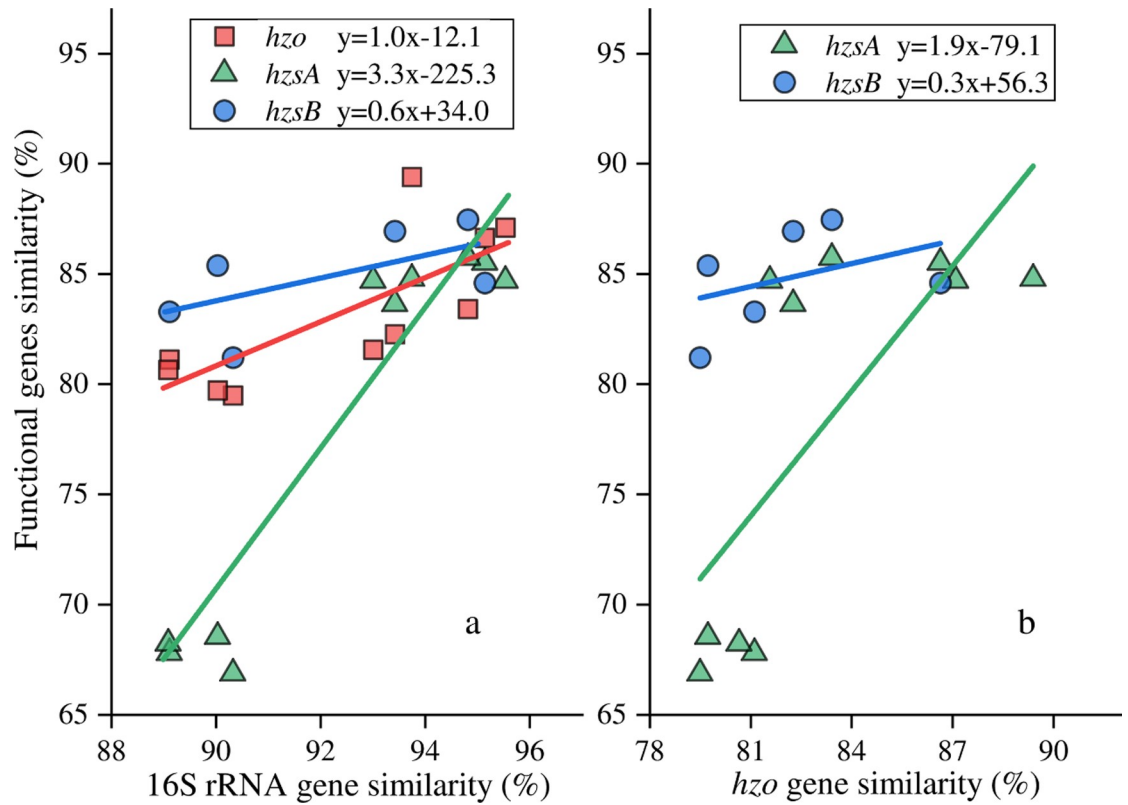
Fig 1. Pairwise nucleotide identity comparisons of functional and 16S rRNA genes of the 5 known genera.
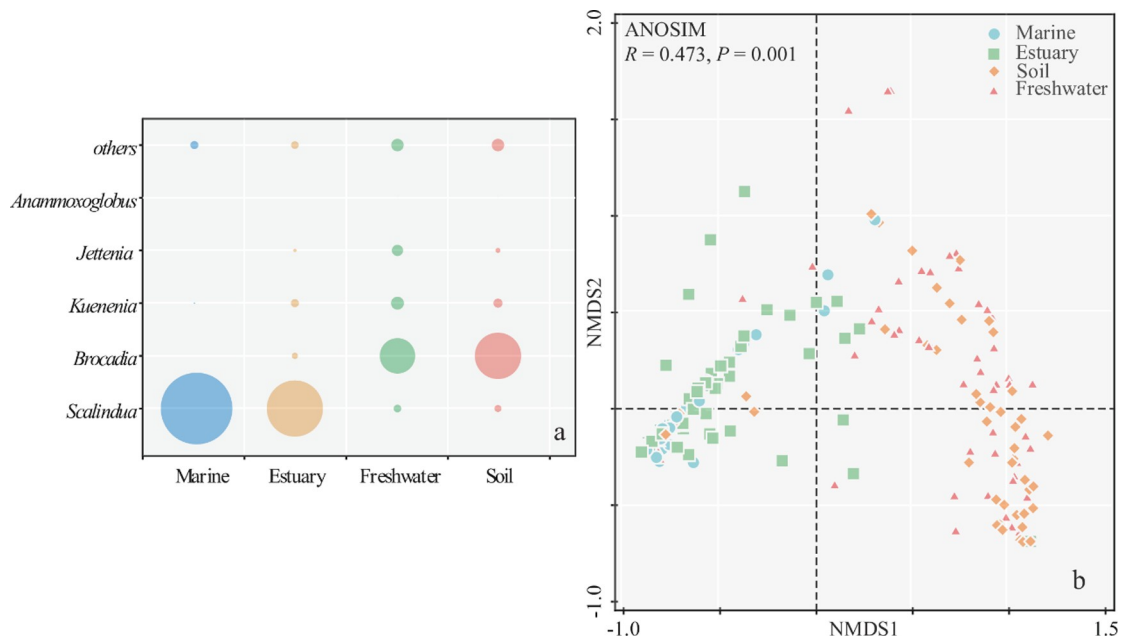
Fig 2. Anammox bacterial community composition in different habitats (a), and NMDS ordination based on Bray-Curtis distance of the sequences affiliation to each genera (b).

Table 1. Number of collected papers with respective marker genes in each habitat.

|	Marine	Estuary	Freshwater	Soil
16S rRNA	34	42	44	31
hzo	17	14	8	4
hzsA	6	2	1	1
hzsB	0	0	4	13

marker gene with 43 studies, and a majority of them were from marine (17) and estuary (14) habitats. Relatively limited studies used the *hzsA* and *hzsB* genes with 10 and 17 studies, respectively.

In marine and estuary habitats, the Ca. *Scalindua* was dominant genus accounting for 88.2% and 69.3% of the total anammox bacteria (Fig 2A). In contrast, the Ca. *Scalindua* only accounted for 9.3% and 8.5% in freshwater and soil environments. Meanwhile, the Ca. *Brocadia* dominated these two environments taking for 43.7% and 57.0% of total anammox bacteria, respectively. Sequences of Ca. *Anammoxoglobus* was observed with relatively higher proportion in soil and freshwater environments than those in marine and estuary. Besides, Ca. *Jettenia* and Ca. *Kuenenia* also took higher proportions in freshwater and soil environments with 13.7% and 15.9%, 6.1% and 11.2%, respectively. The NMDS ordination further showed that the anammox bacterial community composition was distributed according to their habitats, in which the *P* value was 0.001 in ANOSIM comparison. The anammox bacteria in marine and estuary environments were closely distributed, and those of freshwater and soil fell together (Fig 2B).

Community composition revealed by various marker genes

The community composition revealed by the most frequently used two genes were compared in each habitat to show the effect of marker genes (Fig 3A). The anammox bacterial community in marine environment revealed by 16S rRNA and *hzo* genes were very similar, so were those of estuary environment. The only difference was that the Ca. *Jettenia* revealed by *hzo* gene (14.6%) was more abundant than that of 16S rRNA (0.4%) in estuary. Whereas the 16S rRNA gene identified more Ca. *Kuenenia* related sequences (12.1%) than *hzo* gene (0). However, the community composition in freshwater environment revealed by 16S rRNA differed significantly with that of *hzo* gene. The most abundant genus was Ca. *Brocadia* (49.5%)
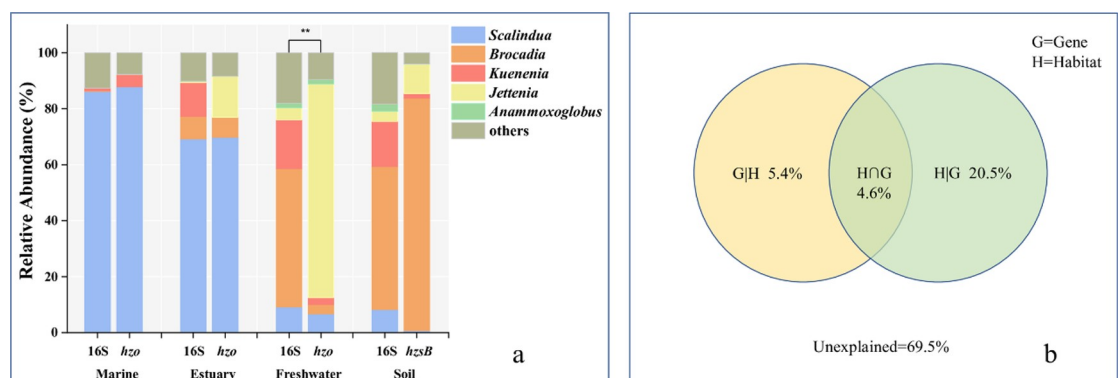
Fig 3. Community composition revealed by the most used two genes in each habitats (a), and Variation Partitioning Analysis (VPA) showing the relative role of habitat and marker gene on the structuring of the anammox bacterial community composition (b).
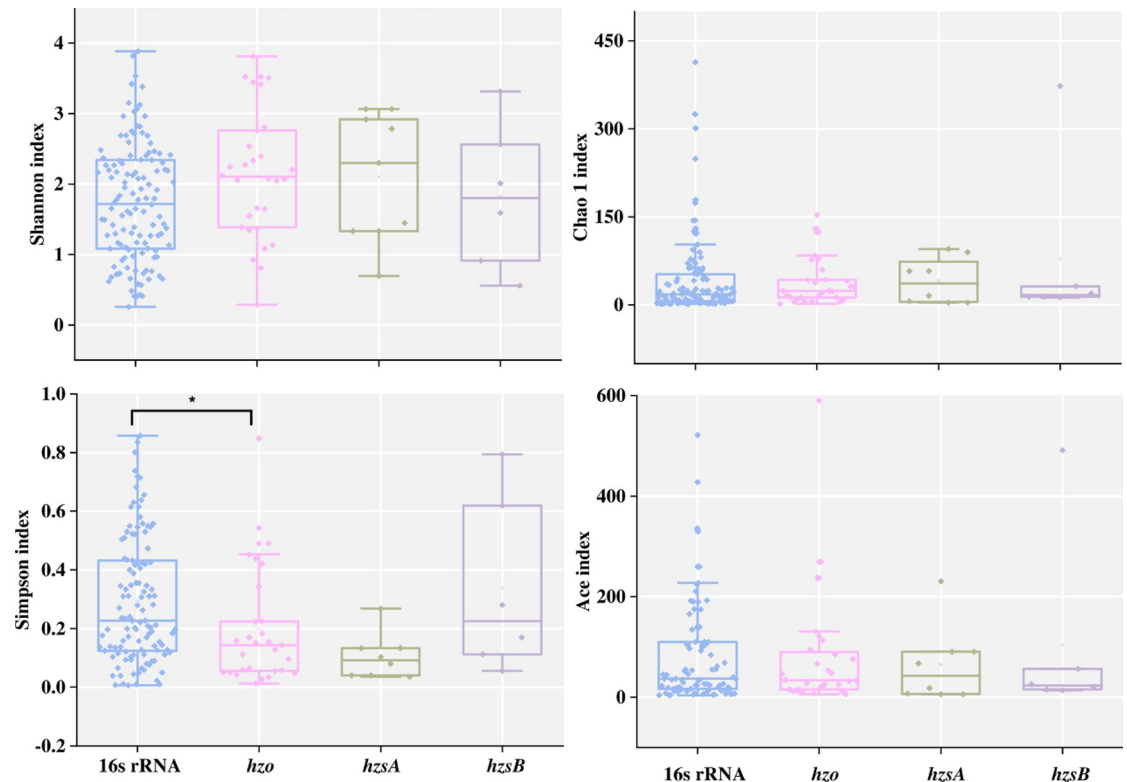
Fig 4. Alpha diversity revealed by the 4 marker genes, significant difference was marked with asterisk (*).

revealed by 16S rRNA gene, followed by Ca. *Kuenenia* (17.4%) and Ca. *Jettenia* (4.3%). The unclassified sequences name by *others* took a considerable proportion of 18.0%. As to *hzo* gene, the Ca. *Jettenia* was predominant taking 76.3% of the total anammox bacterial community. The other four genus took limited proportion which accounting for 14.2% in total. The unclassified group (9.5%) was smaller than that of 16S rRNA gene. No significant difference was observed in soil habitat using 16S rRNA and *hzsB* genes. In the VPA, the factor of sole habitat took a relatively higher explanation with 20.5% of the variance of community composition. The factor of marker gene explained 5.4% of the variance. The cross effect of habitat and marker genes explained 4.6% of the variance.

Alpha diversity revealed by different marker genes

Alpha diversity was assessed by calculating four indexes of Shannon, Simpson, Chao1 and Ace index. The alpha diversity from the 221 studies did not show difference between the 4 habitats (S1 Fig). But the Simpson index revealed by *hzo* gene showed significantly lower values than those of 16S rRNA gene, which suggested higher diversity revealed by *hzo* gene (Fig 4). The other three indexes did not show any difference between the 4 marker genes.

Discussion

PCR amplification is the primary tool to identify anammox bacteria in natural habitats. Multiple marker genes were applied in PCR along with our increasing knowledge of the evolution and metabolism of anammox bacteria [15]. Previous studies showed that different marker

genes might result contrasting community structures [16–18]. However, the bias of marker gene in amplifying anammox bacteria in natural habitats have not been fully investigated.

The genus Ca. *Scalindua* showed the largest difference with other genera presented by all 4 frequently used marker genes, and those amplified by *hzsA* gene showed the highest dissimilarities. In contrast, the dissimilarities between Ca. *Scalindua* and other genera were the lowest presented by 16S rRNA gene. The Ca. *Scalindua* is the dominant genus in marine environment [10]. Therefore, we speculated that the community composition analysis using *hzsA* gene would provide higher resolution when sample from marine was all involved. Besides, the dissimilarities between genera showed comparable values between *hzo* and 16S rRNA genes. It implied a relatively congruent evolutionary relationship of *hzo* and 16S rRNA genes of anammox bacteria.

The 16S rRNA was the most commonly used marker gene for PCR, studies of which were overwhelming (151/221) compared with other marker genes. Besides, most of the cultured anammox bacteria were identified and submitted to public database in form of 16S rRNA gene fragments. There were 17 species of enriched anammox bacteria in GenBank to our statistics [19]. Due to the abundant reference sequences of enriched anammox bacteria, the 16S rRNA gene is a better option in phylogenetic analysis. However, the 16S rRNA gene as a molecular marker is not necessarily related to the physiology of anammox bacteria [12], and the highly conserved nature of the 16S rRNA gene limits its use in revealing the full diversity of anammox bacteria [10]. Analysis of the functional genes involved in specific anammox biochemical reactions could significantly increase the detection efficiency and specificity [20].

The result of anammox bacterial community composition was different using 16S rRNA and *hzo* genes, which revealed more Ca. *Brocadia* and Ca. *Jettenia* in freshwater habitats, respectively. It implied a clear influence of marker genes on the anammox bacterial community composition. The bias of marker genes had been reported that *hzsB* based community analyses showed 3 genera of Ca. *Brocadia*, Ca. *Scalindua* and Ca. *Jettenia* in paddy soil, but the *hzo* gene showed only Ca. *Brocadia* [16]. In marine sediment, the *hzo* gene based diversity analyses revealed 5 genera versus 1 genus by 16S rRNA gene [21]. In a freshwater river, the diversity of anammox bacteria revealed by *hzo* gene was higher than that of 16S rRNA [17]. The contrast of community compositions resulted by different marker genes were mostly caused by the nature 16S rRNA and functional genes which had different detection specificity and efficiency [10]. Besides, the different sources of reference sequences during primers design may lead to bias in amplification efficiency to specific genus [22]. However, no such difference was observed in marine and estuary habitats. It was coincide with previous study that 3 marker genes of 16S rRNA, *hzo*, and *hzsA* all showed that Ca. *Scalindua* was dominant in marine habitats [23]. It was mostly because the diversity on genus level was too low and the genus Ca. *Scalindua* was dominant in these two habitats, that different marker genes resulted similar community composition. It was notable that the marker genes appeared to have no impact on the anammox bacterial community composition in soil, which was coincide with previous study in aquifer sediment using same marker genes [24]. It was probably because that the primers targeting the *hzsB* and 16S rRNA genes have similar specificity and efficiency in amplification. Besides, it also suggested that the role of marker gene in affecting the resulted community composition was limited, as showed in the VPA with low variance from marker gene. Only limited difference was observed on the alpha diversity between marker genes which seemed to be controversial with the community composition analysis. However, we should note that the comparison of alpha diversity was based on sequence level, whereas the comparison of community composition was based on genus level.

In conclusion, we observed the effect of marker gene on the resulted community composition of anammox bacteria in natural habitats, which was significant in freshwater

environments. But the overall effect of marker genes was limited for all natural habitats. The 16S rRNA gene showed advance in phylogenetic analysis and the functional genes revealed more diversities of anammox bacteria on genus level.

Supporting information

S1 Fig. Alpha diversity in the 4 types of habitat.
(DOCX)

S1 Table. Accession number of sequences used in the comparison.
(DOCX)

Acknowledgments

We would like to thank Chunchen Hu, Xiaohan Liu, Yaohao Hu, and Yiben Li for their kind help in experiment and discussion.

Author Contributions

Conceptualization: Yu Wang.

Data curation: Min Cai.

Formal analysis: Min Cai.

Funding acquisition: Yu Wang, Yiguo Hong.

Investigation: Min Cai, Fei Ye, Jiapeng Wu, Qihang Wu.

Methodology: Fei Ye, Jiapeng Wu.

Software: Min Cai.

Supervision: Yu Wang.

Writing – original draft: Min Cai, Yu Wang.

Writing – review & editing: Yiguo Hong.

References

1. Jetten MS, Cirpus I, Kartal B, van Niftrik L, van de Pas-Schoonen KT, Sliekers O, et al. 1994–2004: 10 years of research on the anaerobic oxidation of ammonium. Biochemical Society transactions. 2005; 33 (Pt 1):119–23. https://doi.org/10.1042/BST0330119 PMID: 15667281

2. Schubert CJ, Durisch-Kaiser E, Wehrli B, Thamdrup B, Lam P, Kuypers MM. Anaerobic ammonium oxidation in a tropical freshwater system (Lake Tanganyika). Environmental microbiology. 2006; 8 (10):1857–63. https://doi.org/10.1111/j.1462-2920.2006.01074.x PMID: 16958766

3. Zhu GB, Wang SY, Wang Y, Wang CX, Risgaard-Petersen N, Jetten MS, et al. Anaerobic ammonia oxidation in a fertilized paddy soil. ISME J. 2011; 5(12):1905–12. https://doi.org/10.1038/ismej.2011.63 PMID: 21593796

4. Kuypers MM, Sliekers AO, Lavik G, Schmid M, Jørgensen BB, Kuenen JG, et al. Anaerobic ammonium oxidation by anammox bacteria in the Black Sea. Nature. 2003; 422(6932):608–11. https://doi.org/10.1038/nature01472 PMID: 12686999

5. Kartal B, van Niftrik L, Rattray J, van de Vossenberg JLCM, Schmid MC, Sinninghe Damsté J, et al. Candidatus 'Brocadia fulgida': an autofluorescent anaerobic ammonium oxidizing bacterium. FEMS Microbiology Ecology. 2008; 63(1):46–55. https://doi.org/10.1111/j.1574-6941.2007.00408.x PMID: 18081590

6. Schmid M, Twachtmann U, Klein M, Strous M, Juretschko S, Jetten M, et al. Molecular Evidence for Genus Level Diversity of Bacteria Capable of Catalyzing Anaerobic Ammonium Oxidation. Systematic and Applied Microbiology. 2000; 23(1):93–106. https://doi.org/10.1016/S0723-2020(00)80050-8 PMID: 10879983

7. Kuypers MMM, Sliekers AO, Lavik G, Schmid M, Jørgensen BB, Kuenen JG, et al. Anaerobic ammonium oxidation by anammox bacteria in the Black Sea. Nature. 2003; 422(6932):608–11. https://doi.org/10.1038/nature01472 PMID: 12686999

8. Kartal B, Rattray J, van Niftrik LA, van de Vossenberg J, Schmid MC, Webb RI, et al. Candidatus "Anammoxoglobus propionicus" a new propionate oxidizing species of anaerobic ammonium oxidizing bacteria. Systematic and Applied Microbiology. 2007; 30(1):39–49. https://doi.org/10.1016/j.syapm.2006.03.004 PMID: 16644170

9. Quan Z-X, Rhee S-K, Zuo J-E, Yang Y, Bae J-W, Park JR, et al. Diversity of ammonium-oxidizing bacteria in a granular sludge anaerobic ammonium-oxidizing (anammox) reactor. Environmental microbiology. 2008; 10(11):3130–9. https://doi.org/10.1111/j.1462-2920.2008.01642.x PMID: 18479446

10. Yang Y, Li M, Li H, Li X-Y, Lin J-G, Denecke M, et al. Specific and effective detection of anammox bacteria using PCR primers targeting the 16S rRNA gene and functional genes. Science of The Total Environment. 2020; 734:139387. https://doi.org/10.1016/j.scitotenv.2020.139387 PMID: 32460079

11. Zhou Z, Wei Q, Yang Y, Li M, Gu JD. Practical applications of PCR primers in detection of anammox bacteria effectively from different types of samples. 2018; 102(14):5859–71. https://doi.org/10.1007/s00253-018-9078-2 PMID: 29802476

12. Li M, Gu JD. Advances in methods for detection of anaerobic ammonium oxidizing (anammox) bacteria. Applied microbiology and biotechnology. 2011; 90(4):1241–52. https://doi.org/10.1007/s00253-011-3230-6 PMID: 21476137

13. Acinas SG, Sarma-Rupavtarm R, Klepac-Ceraj V, Polz MF. PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. Applied and environmental microbiology. 2005; 71(12):8966–9. https://doi.org/10.1128/AEM.71.12.8966-8969.2005 PMID: 16332901

14. Ter Braak C, Šmilauer P. Canoco reference manual and user's guide: software of ordination, version 5.0. Ithaca, USA: Microcomputer Power; 2012.

15. Strous M, Pelletier E, Mangenot S, Rattei T, Lehner A, Taylor MW, et al. Deciphering the evolution and metabolism of an anammox bacterium from a community genome. Nature. 2006; 440(7085):790–4. https://doi.org/10.1038/nature04647 PMID: 16598256

16. Gu C, Zhou H, Zhang Q, Zhao Y, Di H, Liang Y. Effects of various fertilization regimes on abundance and activity of anaerobic ammonium oxidation bacteria in rice-wheat cropping systems in China. Sci Total Environ. 2017; 599–600:1064–72. https://doi.org/10.1016/j.scitotenv.2017.04.240 PMID: 28511351

17. Sun W, Xu MY, Wu WM, Guo J, Xia CY, Sun GP, et al. Molecular diversity and distribution of anammox community in sediments of the Dongjiang River, a drinking water source of Hong Kong. J Appl Microbiol. 2014; 116(2):464–76. https://doi.org/10.1111/jam.12367 PMID: 24125160

18. Wang J, Gu JD. Dominance of Candidatus Scalindua species in anammox community revealed in soils with different duration of rice paddy cultivation in Northeast China. Appl Microbiol Biotechnol. 2013; 97 (4):1785–98. https://doi.org/10.1007/s00253-012-4036-x PMID: 22526793

19. Wen XM, Wu JP, Wang Y, Hong YG. Optimization of a method for diversity analysis of anammox bacteria using high-throughput sequencing of 16S rRNA gene amplicon. Submitted. 2020.

20. Kartal B, Geerts W, Jetten MSM. Chapter four—Cultivation, Detection, and Ecophysiology of Anaerobic Ammonium-Oxidizing Bacteria. In: Klotz MG, editor. Methods in Enzymology. 486: Academic Press; 2011. p. 89–108. https://doi.org/10.1016/B978-0-12-381294-0.00004-3 PMID: 21185432

21. Li M, Hong YG, Klotz MG, Gu JD. A comparison of primer sets for detecting 16S rRNA and hydrazine oxidoreductase genes of anaerobic ammonium-oxidizing bacteria in marine sediments. Appl Microbiol Biotechnol. 2010; 86(2):781–90. https://doi.org/10.1007/s00253-009-2361-5 PMID: 20107988

22. Sonthiphand P, Neufeld JD. Evaluating Primers for Profiling Anaerobic Ammonia Oxidizing Bacteria within Freshwater Environments. PLOS ONE. 2013; 8(3):e57242. https://doi.org/10.1371/journal.pone.0057242 PMID: 23505422

23. Han P, Klumper U, Wong A, Li M, Lin JG, Quan Z, et al. Assessment of molecular detection of anaerobic ammonium-oxidizing (anammox) bacteria in different environmental samples using PCR primers based on 16S rRNA and functional genes. Appl Microbiol Biotechnol. 2017; 101(20):7689–702. https://doi.org/10.1007/s00253-017-8502-3 PMID: 28932888

24. Wang SY, Radny D, Huang SB, Zhuang LJ, Zhao SY, Berg M, et al. Nitrogen loss by anaerobic ammonium oxidation in unconfined aquifer soils. Sci Rep. 2017; 7:40173. https://doi.org/10.1038/srep40173 PMID: 28071702